

**Jussi Paavola**

**Klusterin toiminnan tutkiminen Linux-  
ympäristössä**

**Opinnäytetyö  
KESKI-POHJANMAAN AMMATTIKORKEAKOULU  
Tietotekniikan koulutusohjelma, Ylivieska  
Toukokuu 2010**



## TIIVISTELMÄ OPINNÄYTETYÖSTÄ

<b>Yksikkö</b> Ylivieska, Tekniikka	<b>Aika</b> Toukokuu 2010	<b>Tekijä/tekijät</b> Jussi Paavola
<b>Koulutusohjelma</b> Tietotekniikka		
<b>Työn nimi</b> Klusterin toiminnan tutkiminen Linux-ympäristössä		
<b>Työn ohjaaja</b> Hannu Puomio		<b>Sivumäärä</b> 37
<b>Työelämäohjaaja</b> Hannu Puomio		
<p>Työssä on tarkoitus tutkia, että minkälaisia klusterointimahdollisuuksia on tällä hetkellä tarjolla Linux-ympäristöön.</p> <p>Työssä käydään lyhyesti läpi Linux käyttöjärjestelmää ja klusterointia yleisesti. Tarkemmin työssä tutustutaan erilaisiin vaihtoehtoihin, mitä Linuxille on tarjolla klusterointiin. Klusteroinnin yhteydessä, tutkitaan myös hieman pilvitekniikkaa. Työssä rakennetaan myös oma klusteri koulun tietoliikenne-laboratoriossa.</p>		
<b>Asiasanat</b> Eucalyptus, GNU, klusteri, Linux, noodit, serveri, Ubuntu		

**ABSTRACT**

<b>CENTRAL OSTROBOTHNIA UNIVERSITY OF APPLIED SCIENCES</b> Ylivieska, Technology	<b>Date</b> May 2010	<b>Author</b> Jussi Paavola
<b>Degree programme</b> Information Technology		
<b>Name of thesis</b> Exploring clustering in Linux environment		
<b>Instructor</b> Hannu Puomio		<b>Pages</b> 37
<b>Supervisor</b> Hannu Puomio		
<p>The purpose of this thesis was to explore what kinds of clustering programs are available for a Linux platform.</p> <p>In this thesis there is a brief introduction to Linux operating system and Linux clustering in general. The options that there are for Linux clustering are examined more thoroughly. In addition some attention is paid on cloud computing. To explore clustering our own cluster was built in an information technology laboratory in order to test it.</p>		
<b>Key words</b> Cluster, Eucalyptus, GNU, Linux, node, server, Ubuntu		

## LYHENTEET

CPU	Central Processing Unit
DHCP	Dynamic Host Configuration Protocol
DNS	Domain Name System
Gnome	GNU Network Object Model Environment
GNU	GNU's Not Unix
HTML	HyperText Markup Language
IP	Internet Protocol
KDE	K Desktop Environment
LTS	Long Term Support
PCI	Peripheral Component Interconnect
RAID	Redundant Array of Independent Disks
SSH	Secure Shell
UEC	Ubuntu Enterprise Cloud
VM Types	Virtual Machine Types

# SISÄLLYS

TIIVISTELMÄ

ABSTRACT

LYHENTEET

JOHDANTO

1	LINUX.....	2
1.1	Vapaa ohjelmisto GNU.....	2
1.2	Linux -jakelupaketti.....	3
1.3	Ubuntu ja Ubuntu Server -käyttöjärjestelmät.....	3
2	HAJAUTETTU LASKENTA.....	5
2.1	Linux klusterointi.....	5
2.2	Klusterointimallit.....	5
2.3	Klusterointiohjelmat Linuxille.....	6
2.4	Pilvilaskenta (Cloud Computing).....	8
3	JÄRJESTELMÄN ASENNUS JA TESTAUS.....	10
3.1	Asennus.....	10
3.2	Ubuntu Server 9.10 käyttöjärjestelmän asennus.....	10
3.3	Ubuntu Server käyttöjärjestelmän verkkoasetukset.....	13
3.4	Eucalyptus Noodin asennus.....	16
3.5	Eucalyptus Cloud/Cluster/Storage/Walrus servereiden asennus.....	18
4	KLUSTERIN KONFIGUROINTI.....	21
4.1	Klusterin konfigurointi selaimella.....	21
4.2	Levykuvan asennus.....	25
4.3	Levykuvan laittaminen toimintaan.....	27
5	TULOKSET.....	31
6	JOHTOPÄÄTÖKSET.....	33

LÄHTEET

LIITTEET

## **JOHDANTO**

Työssä on tarkoitus tutkia, että mitä erilaisia klusterointimahdollisuuksia Linux- ympäristöön on tällä hetkellä tarjolla. Linux-käyttöjärjestelmiä on monia erilaisia, kuten myös niille tehtyjä klusterointiohjelmia.

Työssä on tutkittu, minkälaisia klusterointiohjelmia on mahdollista käyttää esimerkiksi koulutus tai kotikäytössä.

## **1 LINUX**

Linux on GNU-lisenssin alla toimiva käyttöjärjestelmä, joka on täysin ilmainen. Linuxin kehittäjä Linus Torvalds Unix- järjestelmän pohjalta vuonna 1991. Torvaldsin kirjoittama ydin Linuxiin on vielä tälläkin hetkellä käytössä, tosin sitä on kehitelty kokoajan eteenpäin. Torvalds itse valvoo ytimen kehitystä. Ytimeen voi kuka tahansa tehdä muutoksia.

Ydin eli kerneli (kernel) on tietokoneohjelmiston keskeinen osa, joka mahdollistaa muiden ohjelmistojen toiminnan. Linuxin ytimen päälle voidaan rakentaa erilaisia työpöytäympäristöjä ja ohjelmia. Yleisimmin käytössä olevat työpöytäympäristöt ovat GNU Network Object Model Environment (Gnome) ja K Desktop Environment (KDE).

(Creative Commons 3.0 2010.)

Tänä päivänä Linux on kasvattanut suosiotaan tavallisten kuluttajienkin PC-koneissa, kun sitä aikaisemmin on käytetty lähinnä yliopistoissa ja muissa koulutustarkoituksissa.

### **1.1 Vapaa ohjelmisto GNU**

GNU's Not Unix (GNU) on vuonna 1984 aloitettu projekti, jonka tarkoituksena oli kehittää UNIX- tyylinen käyttöjärjestelmä, joka on täysin ilmainen. GNU:n ja Linuxin ytimen yhdistelmää (GNU/Linux käyttöjärjestelmä) sanotaan lyhyesti Linuxiksi, jota käyttää tänä päivänä miljoonat ihmiset. (Free Software Foundation, Inc. 2010.)

## **1.2 Linux- jakelupaketti**

Jakelupaketti (distribution eli distribuutio) on valmis Linuxin nimen alle rakennettu kokonaisuus. Jakelupaketti sisältää täysin toimivan Linux pohjaisen käyttöjärjestelmän, työpöytäympäristön ja tiettyjä ohjelmia, esimerkiksi internetin käyttöön tarvittavat ohjelmat, toimisto-ohjelmat jne. Kaupallisia jakelupaketteja on myös tarjolla.

(Creative Commons 3.0 2010.)

Suosituimpia ilmaisia jakelupaketteja ovat mm. Ubuntu, Debian Linux, Gentoo, Mandriva, Knoppix, Fedora ja SuSe Linux.

Linux jakelupaketeissa yleisimmät työpöytäympäristöt (graafinen ympäristö) ovat Gnome ja KDE.

## **1.3 Ubuntu ja Ubuntu Server -käyttöjärjestelmät**

Ubuntu on Linux-jakelupaketeista tämän hetken suosituin käyttöjärjestelmä. Se on tehty hyvin käyttäjäystävälliseksi ja sellaisen käyttäjän, joka on aikaisemmin käyttänyt vain Microsoftin Windows -käyttöjärjestelmiä, on melko helppoa siirtyä käyttämään Ubuntu.

Ubuntu on hyvin valmiiksi rakennettu kokonaisuus ja sen mukana tulee normaaliin työpöytäkäyttöön tarvittavat ohjelmat. Mukana on valmiiksi mm. Open Office.org -tekstinkäsittely, taulukkolaskenta -ja esitys-ohjelmat. Muita tarvittavia ohjelmia on mm. Mozilla Firefox nettiselain, Empathy IM messenger/AIM client, Evolution Mail sähköpostiohjema, mediasoitin elokuvien ja videoiden katseluun, musiikinkuunteluohjelma jne. Työpöytäympäristönä Ubuntussa on Gnome, joka on todella



yksinkertainen ja helppo käyttää.

Kaikissa Linux-käyttöjärjestelmissä tulee mukana iptables niminen ohjelma, jolla voi muokata palomuuriasetuksia. Iptablesin sääntöjä voi muokata suoraan terminaalilla/konsolilla tekstipohjaisesti, mutta muokkaamiseen on tehty erilaisia työkaluja. Helppokäyttöisin näistä on varmaankin Firestarter ohjelma, joka pitää asentaa erikseen. Normaalikäytössä riittää kun asentaa Firestarterin, tai ei asenna mitään ohjelmaa. Linux käyttöjärjestelmissä on se hyvä puoli, että sille ei ole tehty viruksia kuin muutama ja haittaohjelmiakaan ei kovin paljoa ole. Linuxiin on hankala hyökätä, joten normaalikäytössä palomuuriasetuksia ei välttämättä tarvitse.

Ubuntu Server käyttöjärjestelmä on tarkoitettu nimensä mukaa erilaisiin palvelimiin. Siinä ei tule mukana ollenkaan työpöytäympäristöä ja asetukset tehdään tekstipohjaisesti. Työpöytäympäristön voi kuitenkin halutessa asentaa ja vaihtoehtoina ovat ainakin Gnome ja KDE. Ubuntun Server käyttöjärjestelmässä tulee mukana melkein kaikki serverille tarkoitetut ominaisuudet. Uusimpaan Ubuntun versioon (9.10) on laitettu mukaan myös klusterointia ja pilvitekniikkaa tukevat ohjelmat.

## **2 HAJAUTETTU LASKENTA**

### **2.1 Linux klusterointi**

Linux klusteroinnin tarkoituksena on saada useampi Linux käyttöjärjestelmä toimimaan yhdessä. Yksi tietokone toimii palvelimena ja ottaa muilta verkon koneilta eli noodeilta esim. prosessorin käyttöä itselleen johonkin raskaaseen tehtävään.

Klusterointia käytetään mm. 3D-kuvan renderöimisessä, tietokannoissa ja palvelimissa. Palvelinklustereita käytetään usein parantamaan järjestelmän suorituskykyä, luotettavuutta ja saavutettavuutta. Useamman keskitehoisen palvelimen muodostama klusteri on myös huomattavasti kustannustehokkaampi kuin yhden superpalvelimen järjestelmä.

### **2.2 Klusterointimallit**

Varsinaiset mallit ovat kahden yksikön järjestelmä ja usean yksikön järjestelmä (esim. ryväslaskenta).  
(Squidoo, LLC 2010.)

Näissä sovellettavat mallit ovat:

- Vektoriprosessointi
- Rinnakkaisprosessointi
- Beowulf

Vektoriprosessoinnissa tehtävät jaetaan useaan osaan ja ne suoritetaan eräänlaisena jonona. Esimerkiksi ensimmäinen kone lukee CD-levyä ja kirjoittaa kiintolevylle raakavedoksen, jonka jälkeen toinen kone kirjoittaa vedoksen toiselle levylle. Vektoriprosessointi eroaa ryväslaskennasta siten, että ryväslaskennassa sama tehtävä ajetaan uudestaan ja uudestaan hieman eri muuttujin joka koneella. Vektoriprosessoinnissa sama tehtävä kulkee jonossa yksiköltä toiselle ja jokainen yksikkö tekee hieman eri tehtävän.

(Creative Commons Attribution/Share-Alike 2010.)

Rinnakkaisprosessoinnissa (ryväslaskenta) kahden noodin mallia voidaan käyttää esim. suojaamaan tärkeä palvelu. Jos ensimmäinen noodin halvaantuu, on toinen noodin heti valmis ottamaan vastaan sen tehtävän. Rinnakkaisprosessoinnissa tehdään myös kuorman jakoa tärkeille verkkopalveluille.

(Creative Commons Attribution/Share-Alike 2010.)

Klusteroinnissa viitataan hajautettuun tiedon käsittelyyn, jota voidaan tehdä myös muulla tavalla kuin laskennallisesti, koska levytila voidaan myös hajauttaa käyttämällä RAID-tekniikkaa. Sama voidaan toteuttaa myös suuremmassa mittakaavassa, tiedon sijainnin kysely tehdään erilliseltä tietokantapalvelimelta joka taas ohjaa hakemaan erilliseltä tietokoneelta kysyttyä tietoa. Tällä tavalla saavutetaan vikasietoinen ja suorituskykyinen järjestelmä.

(Creative Commons Attribution/Share-Alike 2010.)

## **2.3 Klusterointiohjelmat Linuxille**

Linuxille on olemassa paljon eri klusterointiohjelmia.

Vektoriprosessoinnilla toteutettuja ohjelmia ovat mm. :

Corosync Cluster Engine

LifeKeeper

Linux-HA Project

OpenAIS

Red Hat Cluster Suite

Veritas Cluster Server

(Squidoo, LLC 2010.)

Rinnakkaisprosessoinnilla toteutetut ohjelmat:

Keepalived

Linux Virtual Server

Red Hat Cluster Suite

(Squidoo, LLC 2010.)

Beowulf:lla (Compute Clusters/High Performance) toteutetut ohjelmat:

ISLE Cluster Manager

MOSIX

OSCAR

Platform Cluster Manager

Platform HPC Workgroup Manager

Rocks

Scyld Clusterware

xCAT

(Squidoo, LLC 2010.)

Cluster and High Performance File Systems eli tiedostojärjestelmät, jotka on suunniteltu toimimaan klustereissa. Ohjelmia:

OpenAFS

DRBD

GFS

GPFS

Lustre

PVFS

(Squidoo, LLC 2010.)

Lista jatkuu vielä pitkälle, koska Linuxille on myös erilaisia klusteroinnin asennus, seuranta ja hallintaohjelmia.

(Squidoo, LLC 2010.)

Klusteroinnin rakentamiseen on siis paljon erilaisia työkaluja. Vaikea onkin valita seasta sellainen joka olisi kaikkein käytännöllisin. Osa ohjelmista on todella hankala asentaa tai käyttää, mutta kokeneet Linux-käyttäjät ovatkin jo tottuneet, että ohjelmien asentaminen tai käyttäminen vaatii aikaa ja perehtymistä asiaan. Ohjelmat valitaan käyttötarkoituksen mukaan.

Tässä työssä testataan tällä hetkellä uusinta Ubuntun Server -käyttöjärjestelmää, (Ubuntu Server 9.10) joka on ensimmäinen Ubuntun versio, missä on mukana Ubuntu Enterprise Cloud (UEC) niminen klusterointi/pilvi -ohjelma. Tällä tavalla klusterin rakentaminen pitäisi olla melko helppoa.

Työtä tehtäessä testausvaiheen aikana julkaistiin huhtikuussa uusi Ubuntu 10.04 Long Term Support (LTS) -versio. Tässä työssä on käytetty vielä 9.10 versiota.

## **2.4 Pilvilaskenta (Cloud Computing)**

Pilvellä (Cloud) tarkoitetaan web-pohjaista palvelua, joka on ns.

pilvessä. Eli palvelun käyttäjä käyttää esim. jotakin ohjelmaa, joka ei ole asennettu käyttäjän omalle koneelle, vaan ohjelma suoritetaan pilvessä. Näin loppukäyttäjä ei joudu kuluttamaan omalta koneeltaan resursseja ohjelman pyörittämiseen, vaan ohjelma pyörii palveluntarjoajan palvelimella.

Pilvitekniikka on melko uutta tekniikkaa, mutta melkein kaikki virustorjuntaohjelmat alkavat olemaan tänä päivänä pilvessä, joka helpottaa virustietokantojen päivitystä ja loppukäyttäjällä pitäisi olla aina viimeisimmät suojaukset. Virustorjuntaohjelma asennetaan koneelle niin kuin aikaisemminkin, mutta tietokannat ovat palveluntarjoajan palvelimella. Myös virustarkistukset voidaan suorittaa netin kautta palvelimilta.

### **3 JÄRJESTELMÄN ASENNUS JA TESTAUS**

#### **3.1 Asennus**

Työssä oli tarkoitus asentaa Ubuntun Server käyttöjärjestelmä ja sen mukana tuleva klusterointi/pilvi-ohjelma nimeltään Eucalyptus.

Työssä käytettävät tietokoneet olivat koulun tietoliikennelaboratoriossa olevat koneet. Työssä on käytetty kahta konetta. Ensimmäisessä koneessa (serverikone) täytyy olla kaksi verkkokorttia. Toisella verkkokortilla muodostettiin yhteys koulun verkon kautta nettiin ja toinen verkkokortti yhdistettiin ristikaapelilla toisen koneen (noodikone) verkkokorttiin. Koneet olivat kokoonpanoltaan identtisiä:

Fujitsu Siemens Scaleo T D1826

Proessori: Intel Pentium 4 CPU 3.40GHz

Näytönohjain: ATI Radeon X600/X550 Series

DVD-asema: HL-DL-ST DVD-RW GWA-4160B

Muistit: 2 x Samsung 512MB DDR

Verkkokortti: 1: ADMtek Comet (eth0) 2: SiS 900 PCI Fast Ethernet (eth1)

Kiintolevyt: Maxtor 40GB ja 60GB

#### **3.2 Ubuntu Server 9.10 käyttöjärjestelmän asennus**

Ubuntun asennus aloitetaan menemällä osoitteeseen

<http://www.ubuntu.com/> ja ladataan sieltä Ubuntun Server-

käyttöjärjestelmä (32-bittinen versio). Tässä työssä käytettiin 32-bittistä versiota. Kun Image on ladattu, poltetaan se mieluiten CD-RW-levylle. Laitetaan CD sisään ja käynnistetään tietokone uudelleen.

Asennusohjelma käynnistyy ja aloitusvalikko näyttää seuraavalta:



KUVIO 1. Ubuntun asennus (Canonical Ltd 2010.)

Valitaan "Install Ubuntu Enterprise Cloud". Sen jälkeen asennus tutkii laitteiston ja kysyy mitä asennetaan.

(Canonical Ltd 2010.)



KUVIO 2. Valintaruutu (Canonical Ltd 2010.)

Asennusohjelma kysyy, että asennetaanko "Cluster" vai "Node". Valitaan



"Cluster". Sen jälkeen asennus kysyy normaaleita asioita, mitä käyttöjärjestelmän asennuksessa yleensäkin, mm. klusterin nimi, koneen nimi, käyttäjän nimi, salasana, verkkoasetukset, kiintolevyn osiointi ym. Jossakin vaiheessa asennus kysyy, mitä IP-osoite- aluetta klusterille käytetään. Valitaan esimerkiksi se, mitä asennus valmiiksi ehdottaa, eli 192.168.1.100 – 192.168.1.192. Serverille (eli klusterille) annetaan IP-osoitteeksi 192.168.1.101 (Canonical Ltd 2010). Kun asennus on valmis asennusohjelma antaa CD-levyn ulos asemasta ja kone käynnistetään uudestaan.

Noodi-koneelle (eli node) asennetaan käyttöjärjestelmä samalla tavalla, mutta alussa valitaan "Node" eikä "Cluster" (Canonical Ltd 2010). Noodille annetaan IP-osoitteeksi 192.168.1.102. Noodin asennuksessa asennus- ohjelma valittaa, että tietokoneen prosessori ei tue virtualisointia ja saattaa hidastaa klusterin toimintaa. Jatketaan asennusta silti normaalisti.

Kun molempiin koneisiin on asennettu Ubuntun Server -käyttöjärjestelmä, aloitetaan konfigurointi. Server käyttöjärjestelmässähän ei tule olleenkaan mukana graafista käyttöliittymää, joten konfiguroinnit pitää tehdä tekstipohjaisesti. Työn tekemisen helpottamiseksi asennetaan kuitenkin kumpaankin koneeseen Gnome-työpöydät. Ensin täytyy saada nettiyhteys toimimaan, jotta Gnomen voi asentaa. Kun nettiyhteys on saatu määritettyä serveri-koneeseen (joudumme käyttämään koulun välityspalvelinta eli proxyä: 195.148.65.5:800), päivitetään käyttöjärjestelmään uusimmat päivitykset (paketit). Uusimpien päivitysten asennus tapahtuu seuraavasti.

Päivitetään pakettilistat:

`sudo apt-get update` (*sudo komennon edessä tarkoittaa sitä, että*

*komento suoritetaan root-käyttäjänä, eli järjestelmänvalvojan oikeuksilla.)*

Asennetaan uusimmat paketit:

```
sudo apt-get dist-upgrade
```

Kun uusimmat paketit on asennettu, asennetaan Gnome työpöytä:

```
sudo apt-get install ubuntu-desktop
```

Kun työpöytä on asennettu käynnistetään se komennolla:

```
startx
```

### **3.3 Ubuntu Server käyttöjärjestelmän verkkoasetukset**

Nyt on serverikoneessa asennettuna työpöytä. Noodikoneeseen työpöydän asentaminen on vähän hankalampaa, koska se ei ole suoraan netissä. Jotta nettiyhteys saadaan toimimaan myös noodikoneella, täytyy tehdä lisää konfigurointeja. Yhteys muodostetaan serverikoneen kautta. Noodikoneessa on vain yksi verkkokortti, joka on kiinni serverikoneen toisessa verkkokortissa. Eli yhteys pitää muodostaa serverikoneen kautta.

Laitetaan molemmille koneille IP-osoitteet kuntoon. Voimassa olevat IP-osoitteet näkee komennolla `ifconfig`. Serverikoneen `ifconfig`-tiedoston pitäisi näyttää, että koneessa on kaksi verkkokorttia `eth0` ja `eth1`. Noodikoneessa vain yksi eli `eth0`.

Käytetään terminaalissa/konsolissa `nano` nimistä tekstieditoria ja mennään seuraavaan hakemistoon:

```
sudo nano /etc/network/interfaces
```

Serverikoneen /etc/network/interfaces -tiedosto:

```
# This file describes the network interfaces available on  
your system
```

```
# and how to activate them. For more information, see  
interfaces(5).
```

```
# The loopback network interface
```

```
auto lo
```

```
iface lo inet loopback
```

```
# The primary network interface
```

```
auto eth0
```

```
iface eth0 inet dhcp
```

```
auto eth1
```

```
iface eth1 inet static
```

```
address 192.168.1.101
```

```
netmask 255.255.255.0
```

Noodin /etc/network/interfaces -tiedosto:

```
auto lo
```

```
iface lo inet loopback
```

```
auto br0
```

```
iface br0 inet static
```

```
    address 192.168.1.102
```

```
    network 192.168.1.100
```

```
    netmask 255.255.255.0
```

```
    gateway 192.168.1.101
```

```
    broadcast 192.168.1.255
```

```
bridge_ports eth0
bridge_fd 9
bridge_hello 2
bridge_maxage 12
```

Laitetaan kummallekin koneelle niiden nimet ja IP-osoitteet seuraavaan tiedostoon:

```
sudo nano /etc/hosts
```

Serverin /etc/hosts -tiedosto:

```
127.0.0.1      localhost
127.0.1.1      ykay-c23501.edu.local  ykay-c23501
192.168.1.101  ykay-c23501
192.168.1.102  noodi
```

Noodin /etc/hosts -tiedosto:

```
127.0.0.1      localhost
127.0.1.1      noodi
192.168.1.101  ykay-c23501
192.168.1.102  noodi
```

Iptablesin sääntöihin täytyy tehdä muutoksia serverikoneella, jotta se osaa jakaa nettiyhteyden eteenpäin noodikoneelle. Jakaminen tehdään seuraavasti:

Mennään tekstieditorilla seuraavaan tiedostoon serverikoneella:

```
sudo nano /etc/rc.local
```

Lisätään sinne muutama rivi, eli nettiyhteyden jakamissäännöt:

```
# By default this script does nothing.
```

```
iptables -A FORWARD -i eth1 -o eth0 -s 192.168.1.0/16 -m
```

```
conntrack --ctstate NE$
iptables -A FORWARD -m conntrack --ctstate
ESTABLISHED,RELATED -j ACCEPT
iptables -A POSTROUTING -t nat -j MASQUERADE
(RodGer 2010.)
```

Käynnistetään molemmissa koneissa verkko vielä uudestaan komennolla:

```
sudo /etc/init.d/networking restart
```

Nyt voidaan noodikoneeseenkin asentaa uusimmat päivitykset ja Gnome työpöytä.

### 3.4 Eucalyptus Noodin asennus

Seuraavaksi ohjeessa käsketään asentaa noodin paketit.

```
sudo apt-get install eucalyptus-nc
```

(Kirkland 2010.)

Kyseinen paketti on jo asennettu ja on uusin versio. Paketti on asennettu käyttöjärjestelmän asennuksen yhteydessä. Seuraavaksi pitää noodille konfiguroida bridge eli silta. Kyseiset asetukset tehdään /etc/network/interfaces -tiedostoon. Nämä konfiguroinnit on tehty jo aikaisemmin kohdassa 3.3.

Seuraavaan tiedostoon /etc/eucalyptus/eucalyptus.conf käydään antamassa sillalle oikea nimi, joka on "br0". Käynnistetään noodiserveri uudestaan komennolla:

```
sudo /etc/init.d/eucalyptus-nc restart
```

(Kirkland 2010.)

Noodille täytyy antaa verkkoavaimet. Se tapahtuu seuraavasti. Tehdään Eucalyptus niminen käyttäjä väliaikaisesti, jotta voidaan asettaa Secure Shell (SSH) verkkoavaimet:

```
sudo passwd eucalyptus
```

(Kirkland 2010.)

Sitten serverikoneessa laitetaan konsoliin:

```
sudo -u eucalyptus ssh-copy-id -i ~eucalyptus /.ssh  
/id_rsa.pub eucalyptus@<IP_OF_NODE>
```

(Kirkland 2010.)

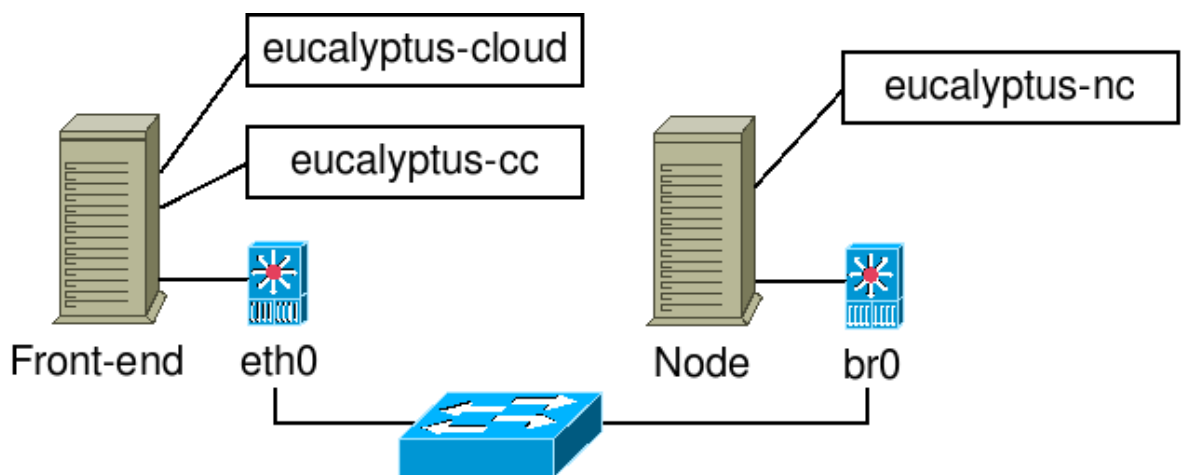
Mennään serverillä noodikoneelle SSH-yhteydellä ja tallennetaan verkkoavaimet. Sen jälkeen noodilta voi poistaa luodut tunnukset:

```
sudo passwd -d eucalyptus
```

(Kirkland 2010.)

Nyt pitäisi noodin asetukset olla kunnossa ja voidaan kokeilla, että yhteys toimii koneiden välillä. Yhteys on jo testattu, koska SSH:lla saatiin yhteys noodiin. Voidaan kuitenkin vielä varmuuden vuoksi pingata (eli lähettää yksittäisiä paketteja) koneelta toiselle komennolla `ping <IP-osoite>`. Molemmat koneet pingaavat toisiaan, eli yhteys toimii.

Tehtyjen konfigurointien jälkeen, asennuksen pitäisi näyttää suurinpiirtein tältä.



KUVIO 3. Noodin asennus (Kirkland 2010.)

### 3.5 Eucalyptus Cloud/Cluster/Storage/Walrus serverien asennus

Seuraavat paketit täytyy olla asennettuna, jotta konfiguroinnit toimivat:

```
sudo apt-get install eucalyptus-cloud eucalyptus-cc
eucalyptus-walrus eucalyptus-sc
```

(Kirkland 2010.)

Kyseiset paketit on jo asennettu ja niistä on uusimmat versiot. Seuraavaksi Cloud/Cluster/Storage/Walrus serverit pitää rekisteröidä. Rekisteröinnin pitäisi tapahtua automaattisesti, mutta jos näin ei ole, niin ne rekisteröidään seuraavasti:

```
sudo start eucalyptus-cc-registration jne.
```

(Kirkland 2010.)

Rekisteröinti on tapahtunut automaattisesti, koska kaikki serverit ovat samalla koneella. Rekisteröinnin onnistumisen voi tarkistaa seuraavasti:

```
klusteri@ykay-c23501:~$ tail -n1
/var/log/eucalyptus/*registration.log
```

```
==> /var/log/eucalyptus/cc-registration.log <==
SUCCESS: new cluster 'KLUSTERI' on host '10.65.0.9'
successfully registered.
```

```
==> /var/log/eucalyptus/sc-registration.log <==
SUCCESS: new SC for cluster 'KLUSTERI' on host '10.65.0.9'
successfully registered.
```

```
==> /var/log/eucalyptus/walrus-registration.log <==
SUCCESS: new walrus on host '10.65.0.9' successfully
registered.
```

Rekisteröinti on tehty onnistuneesti.

Noodi täytyy myös rekisteröidä. Serverikoneelta etsitään verkossa olevat noodit seuraavasti:

```
sudo euca_conf --no-rsync --discover-nodes
(Kirkland 2010.)
```

Noodien etsinnässä tulee ongelmia, koska kyseinen komento jää odottamaan välityspalvelinta, vaikka ei pitäisi:

```
klusteri@ykay-c23501:~$ sudo euca_conf --no-rsync --discover-
nodes
```

WARNING: Unhandled message:

```
interface=org.freedesktop.DBus.Introspectable, path=/,
member=Introspect
```

```
--2010-04-21 13:28:40--
```

```
http://127.0.0.1:8774/axis2/services/
```

```
Connecting to 195.148.65.5:800... connected.
```

```
Proxy request sent, awaiting response... 403 Forbidden
```



2010-04-21 13:28:40 ERROR 403: Forbidden.

ERROR: you need to be on the CC host and the CC needs to be running.

Lopulta noodit löytyy, kun molemmilta koneilta on otettu välityspalvelin-asetukset pois.

```
klusteri@ykay-c23501:~$ sudo euca_conf --no-rsync
--discover-nodes
```

```
[sudo] password for klusteri:
```

WARNING: Unhandled message:

```
interface=org.freedesktop.DBus.Introspectable, path=/,
member=Introspect
```

```
New node found on 192.168.1.102; add it? [Yn] y
```

```
--2010-04-28 10:32:43--
```

```
http://127.0.0.1:8774/axis2/services/
```

```
Connecting to 127.0.0.1:8774... connected.
```

```
HTTP request sent, awaiting response... 200 OK
```

```
Length: 730 [text/html]
```

```
Saving to: `STDOUT'
```

```
100%[=====>] 730
```

```
--.-K/s in 0s
```

```
2010-04-28 10:32:43 (73.8 MB/s) - '-' saved [730/730]
```

INFO: We expect all nodes to have eucalyptus installed in //var/lib/eucalyptus/keys for key synchronization.

Trying scp to sync keys to:

```
eucalyptus@192.168.1.102://var/lib/eucalyptus/keys/...
done.
```

Eli nyt on noodit myös rekisteröity.

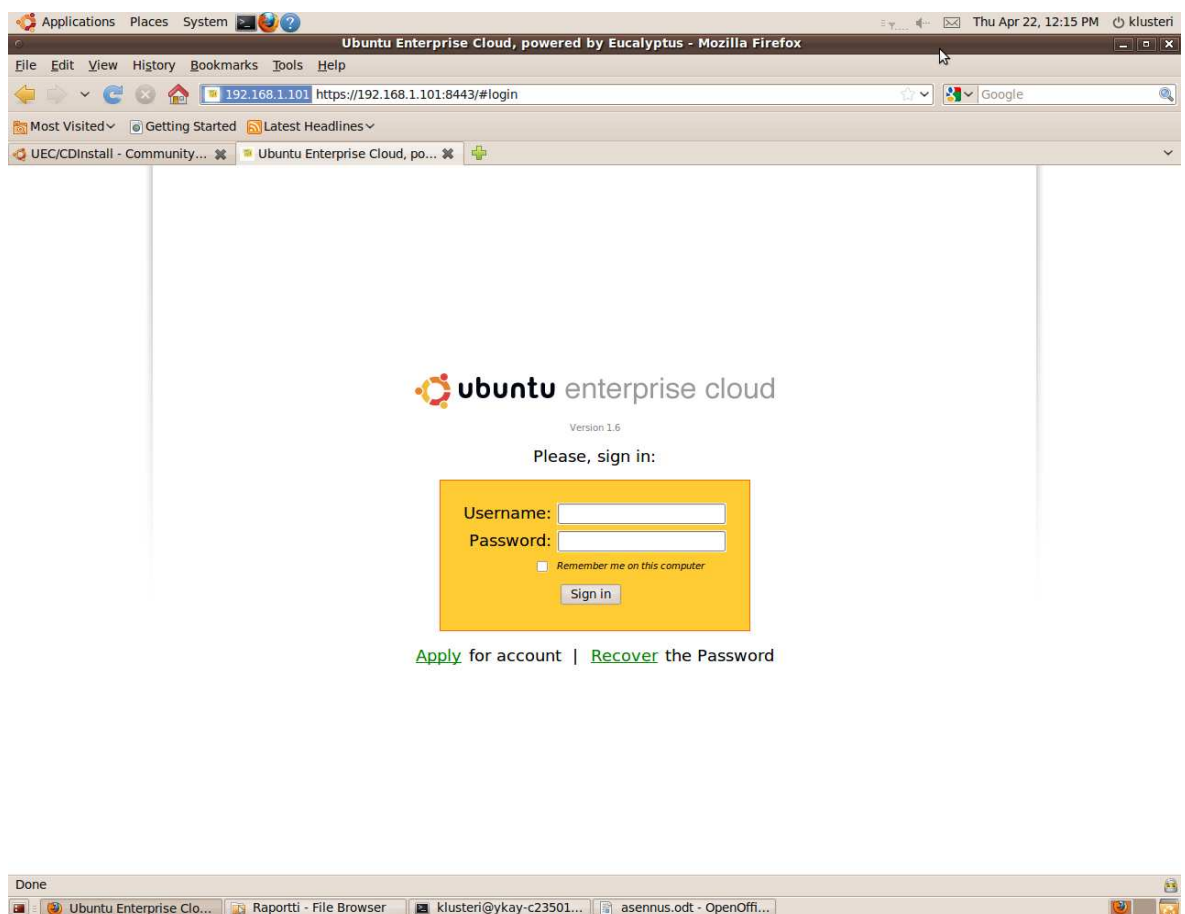
## 4 KLUSTERIN KONFIGUROINTI

### 4.1 Klusterin konfigurointi selaimella

Klusterissa tulee mukana selaimella hallittava Ubuntu Enterprise Cloud. Hallintaan pääsee käsiksi kirjoittamalla selaimen (Mozilla Firefox) osoitekenttään seuraavasti.

`https://<cloud-controller-ip-address>:8443/`  
(Kirkland 2010.)

Tässä tapauksessa IP-osoitteksi kirjoitetaan serverin IP eli 192.168.1.101. Kirjautumisikkuna näyttää seuraavalta.



KUVIO 4. Kirjautuminen

Ensimmäisellä kirjautumiskerralla käytetään käyttäjänimenä "admin" ja salasana on myös "admin".

(Kirkland 2010.)

Applications Places System Thu Apr 22, 10:56 AM klusteri

Ubuntu Enterprise Cloud, powered by Eucalyptus - Mozilla Firefox

File Edit View History Bookmarks Tools Help

192.168.1.101 https://192.168.1.101:8443/#login

Most Visited Getting Started Latest Headlines

UEC/CDInstall - Community... UEC/PackageInstall9.10 - C... Ubuntu Enterprise Cloud, po...

**ubuntu enterprise cloud**

First-time Configuration

Please, supply the following parameters to activate your Eucalyptus installation:

Administrator's new password:

The password, again:

Administrator's email address:

User signup requests will be sent to this address and messages from Eucalyptus will have this address in the From: field. (If you want to change this behavior, edit the appropriate values in the eucalyptus-web.properties file of your Eucalyptus installation. We still need an email address now, though.)

Cloud Host:

We made a guess about the external IP of the machine running the Cloud Controller. Please, make sure that it is correct, as this IP will be embedded in the credentials generated by Eucalyptus. Although it can be changed later, doing so will require that you perform extra tasks and may cause existing users from being unable to access the system.

Submit

Done

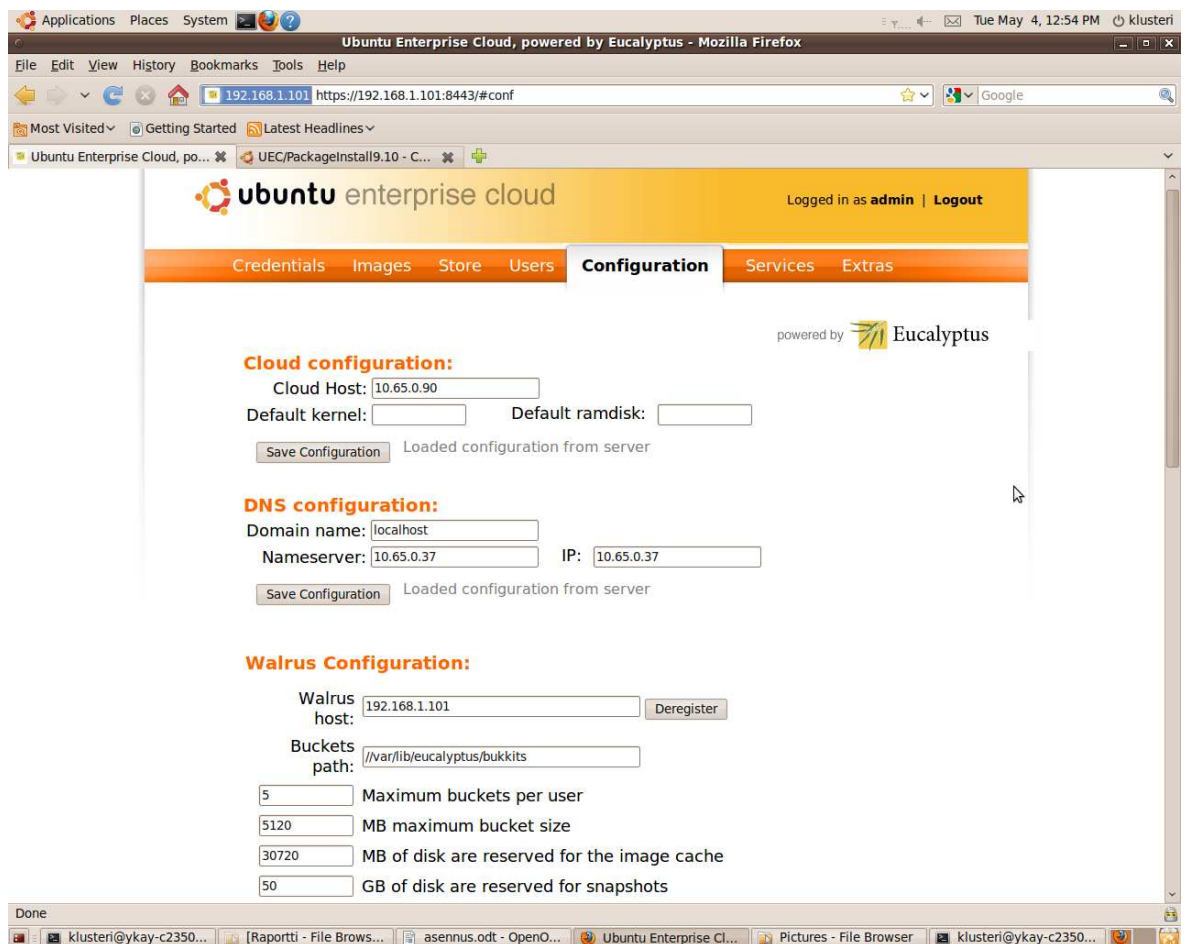
Ubuntu Enterprise Clo... [\*Unsaved Document ... klusteri@ykey-c23501...

## KUVIO 5. Ensimmäinen sisäänkirjaus

Kun ensimmäisen kerran ollaan kirjaututtu sisään, täytyy vaihtaa heti salasana. Annetaan myös toimiva sähköpostiosoite, joka on adminin eli pääkäyttäjän sähköposti.

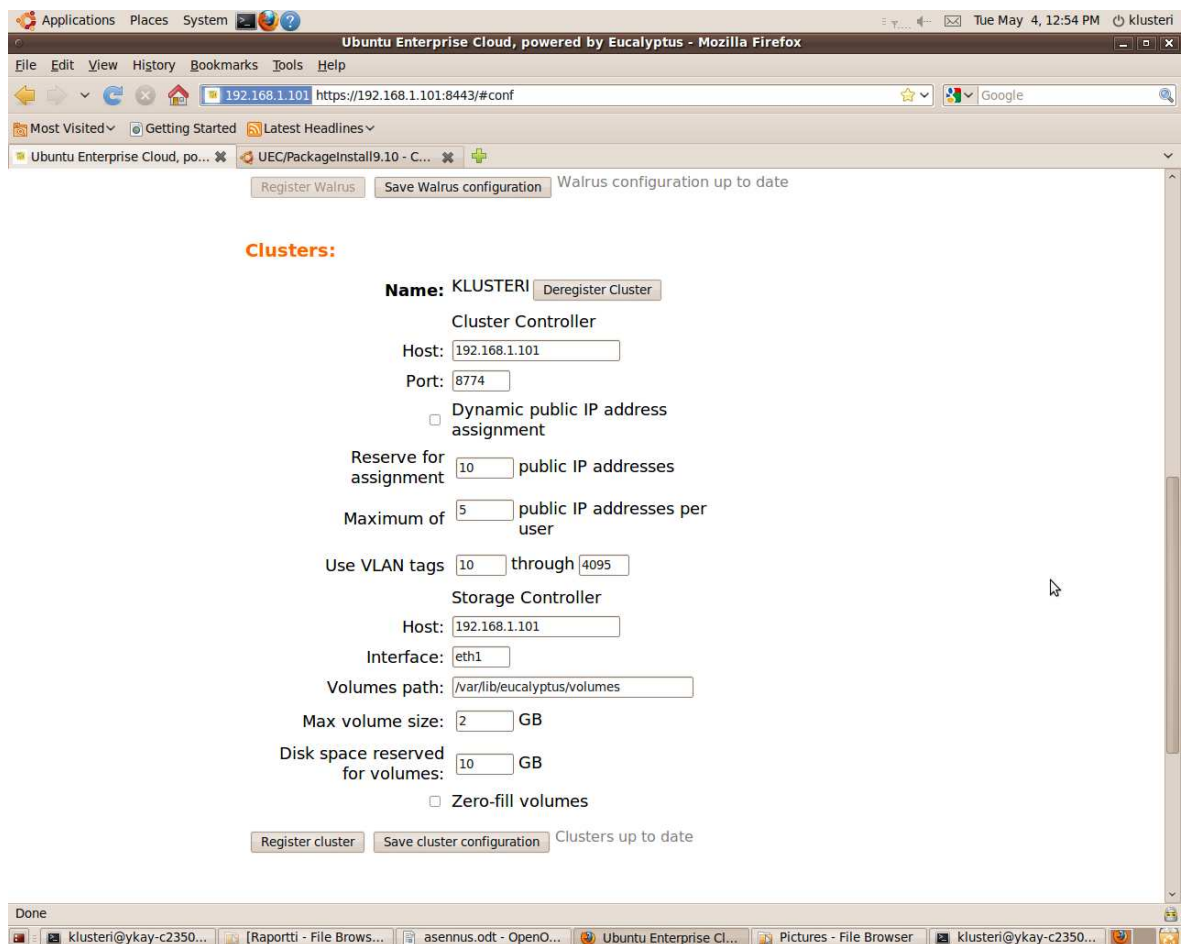
(Kirkland 2010.)

Seuraavaksi katsotaan klusterin konfiguroinnit kuntoon kohdassa "Configuration".



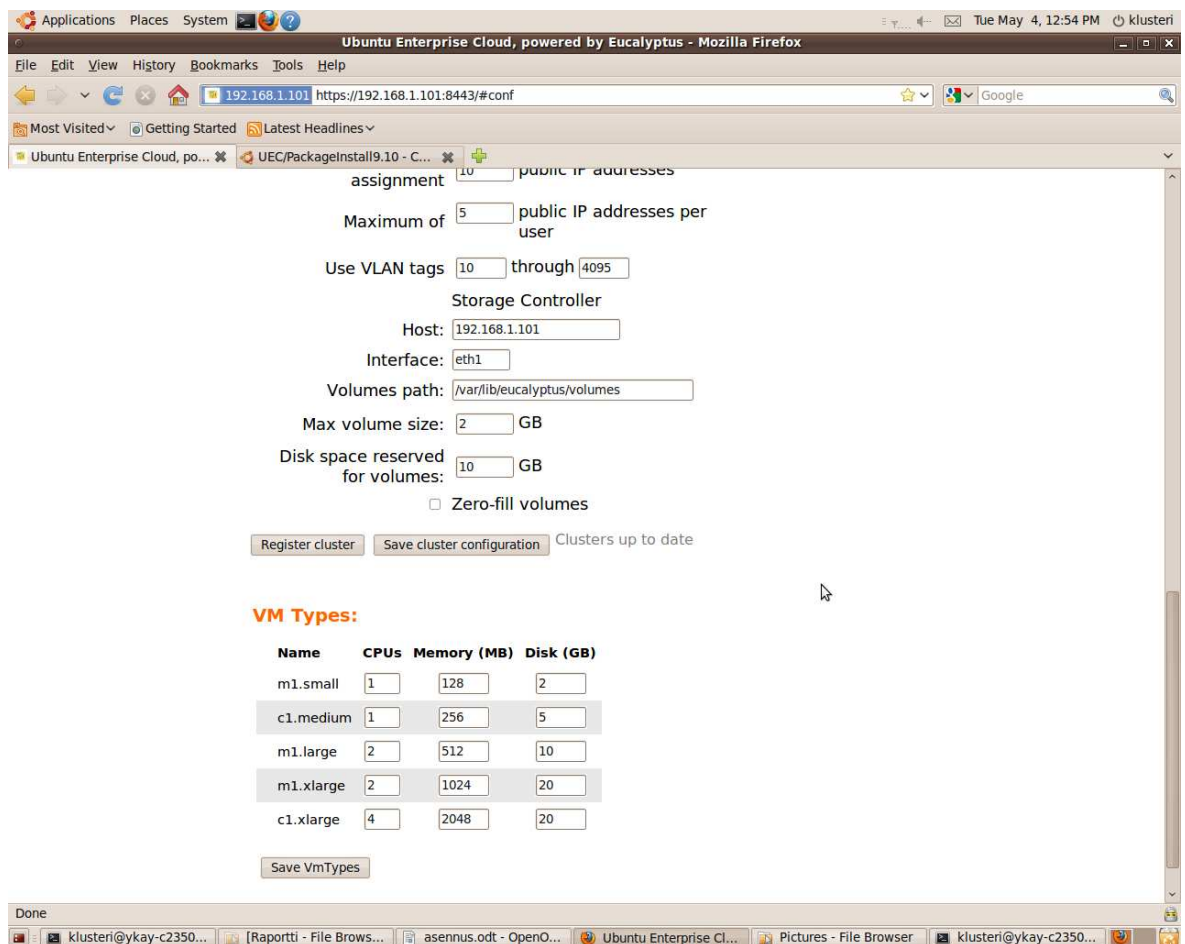
## KUVIO 6. Konfigurointi

”Cloud configuration” kohdassa on serverikoneen julkinen IP-osoite, joka on tässä tapauksessa 10.65.0.90 ja tämän osoite tulee DHCP:n kautta. ”DNS configuration” kohdassa on myös valmiiksi osoite, joten siihen ei tarvitse koskea. ”Cloud configuration” kohdassa kannattaa tarkistaa, että osoite on sama, minkä ifconfig komento konsolissa näyttää eth0 verkkokortille. ”Walrus configuration” kohdassa on sitten serverin sisäverkon IP-osoite eli 192.168.1.101. Kun sivua vieritetään alaspäin, tulee lisää asetuksia.



## KUVIO 7. Konfigurointi

"Clusters" kohdassa ilmoitetaan serverillä olevat klusterit. Kuten kuvasta näkyy, meillä on KLUSTERI niminen klusteri jonka IP-osoite on 192.168.1.101. "Storage controller" kohdassa tarkistetaan, että IP on sama kuin edellisessä kohdassa. Syytä on myös tarkistaa, että "Interface" kohdassa on "eth1", eikä "eth0". Serverikoneessa olevan kiintolevyn pienen koon vuoksi, joudumme pienentämään Storage controllerille varaamaa levytilaa. Muutetaan "Max volume size" kohtaan 2 GB ja "Disk space reserved for volumes" kohtaan 10 GB. Kun sivua rullataan alaspäin, tulee vielä hieman asetuksia.

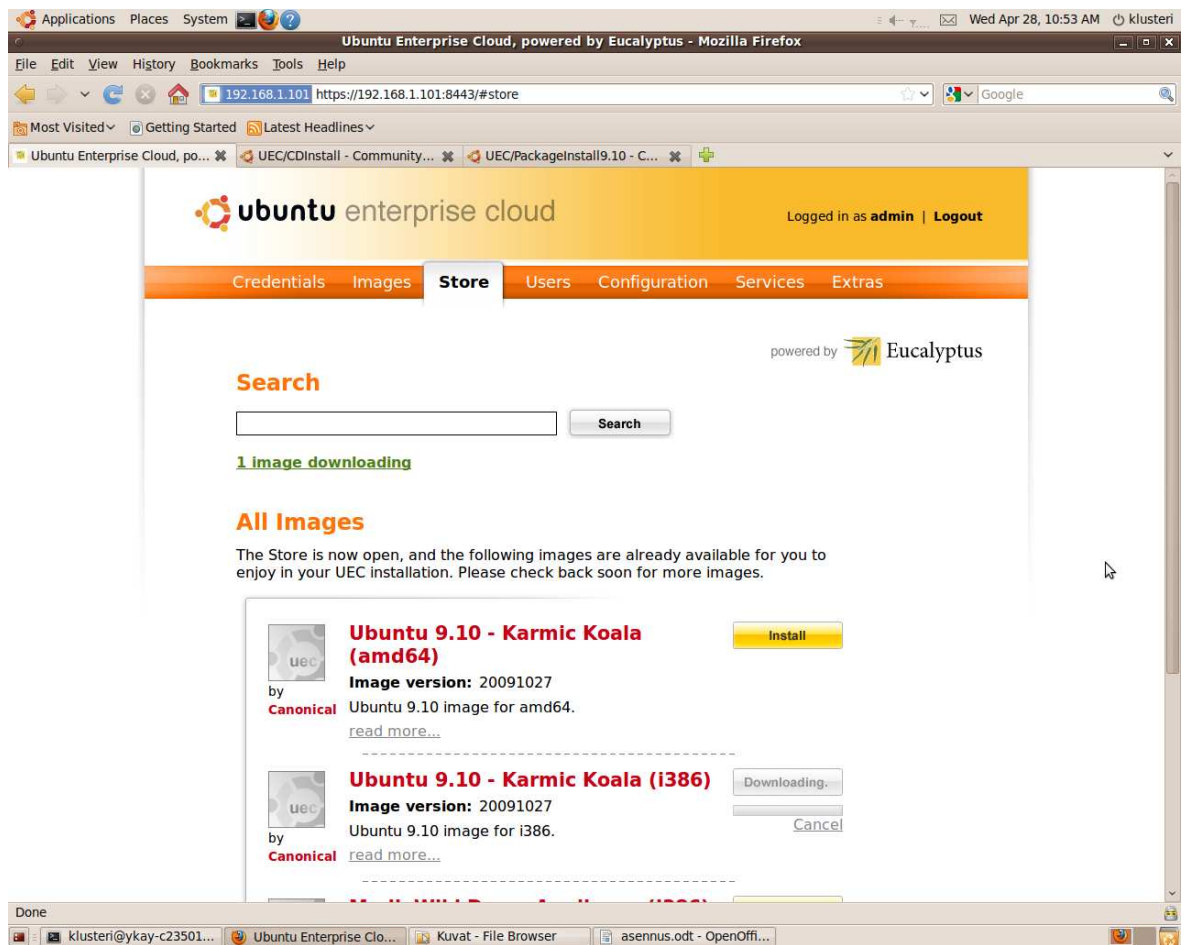


## KUVIO 8. Konfigurointi

”VM Types” kohdassa on valmiiksi laitettu arvot, mitä virtuaalinen image käyttää. Näihin asetuksiin ei tarvitse koskea.

## 4.2 Levykuvan asennus

Levykuva eli image pitää asentaa serverikoneelle, jotta klusteri toimii. Levykuvan asennuksessa ilmeni ongelmia, levykuvaa ei saatu ladattua.

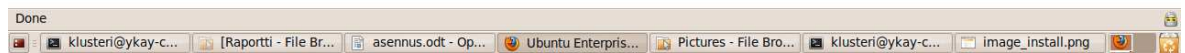
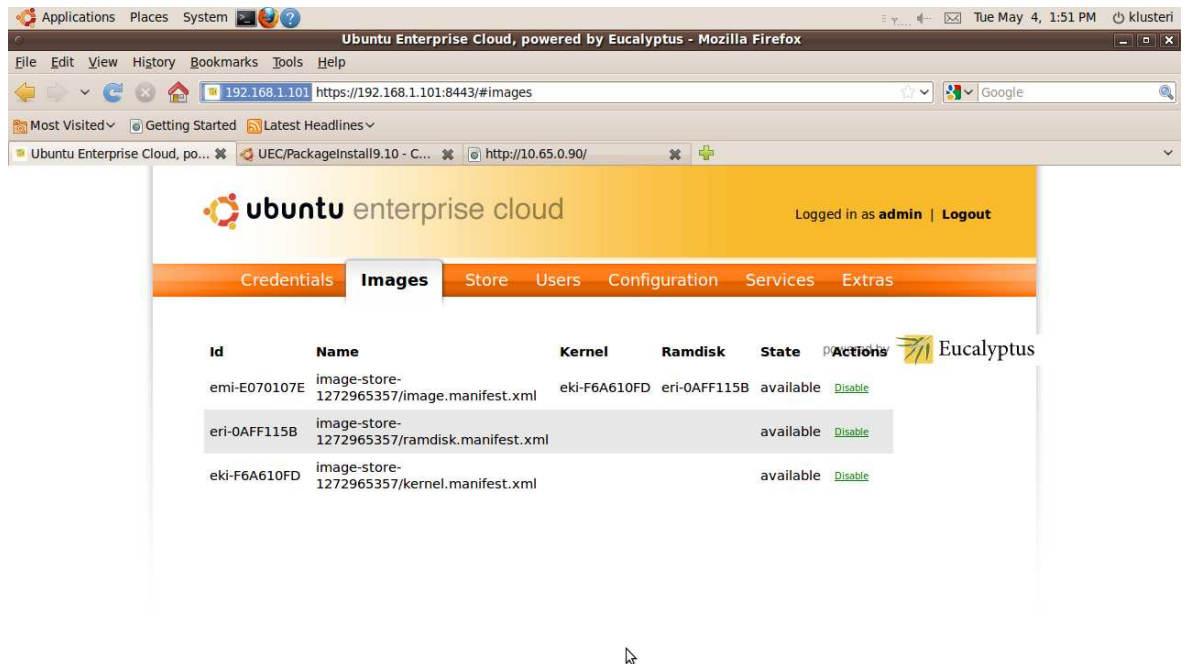


## KUVIO 9. Imagen asennus

Painetaan "install" Ubuntu 9.10 – Karmic Koala (i386) imagen kohdalla. Sen jälkeen "install" napin tilalla lukee "Downloading". Hetken päästä tulee seuraavanlainen virheilmoitus:  
Error 28: connect() timed out!

Tämän ongelman selvittämiseen meni kauan aikaa, mutta lopulta Eucalyptuksen foorumeilta löytyi ratkaisu. Poistettiin paketti nimeltään python-image-store-proxy. Sen jälkeen päivitettiin pakettilistat ja asennettiin kyseinen paketti uudestaan. Eli ongelmana oli taas välityspalvelin. Kun paketti on asennettu uudestaan, imagen lataus onnistuu ongelmitta.

Nyt levykuva on asennettu onnistuneesti.



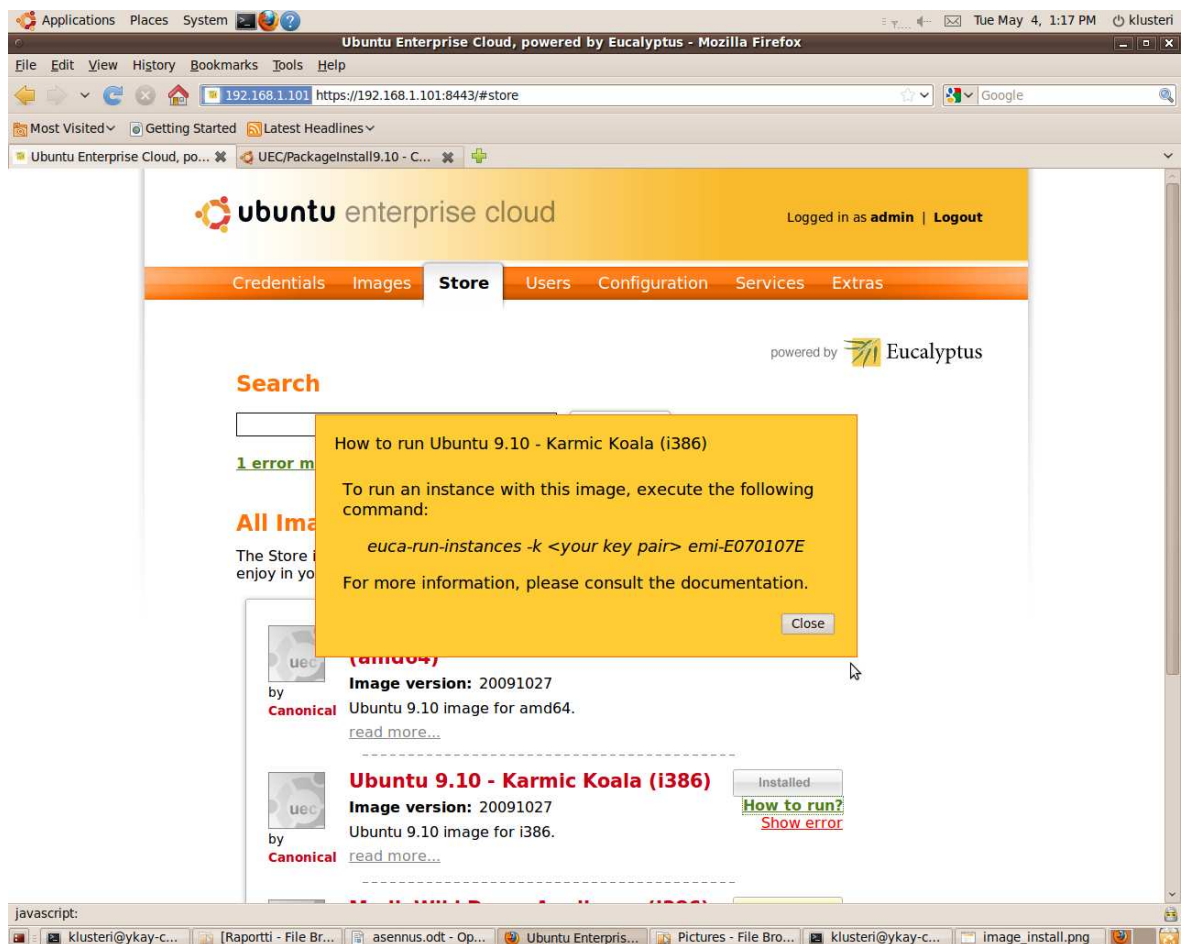
## KUVIO 10. Imaget

Kohdassa "Images" näemme vasta asennetun levykuvan.

### 4.3 Levykuvan toimintaan laittaminen

Levykuvan toimintaan laittaminen tapahtuu seuraavasti.





KUVIO 11. Levykuvan ohje

Menemme kohtaan "Store" ja näemme, että olemme asentaneet Ubuntu 9.10 – Karmic Koala (i386) imagen. Kun painaa tekstiä "How to run?" ilmestyy yllä kuvassa näkyvä ikkuna. Ikkunassa lukee ohje, miten imagen saa toimimaan. Ohje käskee laittaa verkkoavaimen, jotta image toimisi. Verkkoavaimet saa ladattua "Credentials" kohdasta zip-tiedostona. Kun zip-paketti on purettu laitetaan kuvassa näkyvään komentoon polku, mihin paketti purettiin:

```
euca-run-instances -k /home/klusteri/euca2-admin-x509/euca2-admin-3adbb3d2-pk.pem emi-E070107E
```

Sitten tulee seuraavanlainen varoitus:

EC2\_ACCESS\_KEY environment variable must be set.

Tämä varoitus tarkoittaa sitä, että EC2 työkaluja ei ole otettu käyttöön. Otetaan ne käyttöön komennolla:

```
source .eucarc
```

Tämän jälkeen kokeillaan uudestaan ottaa image käyttöön. `euca-run-instances` komennon jälkeen konsolissa ei lue mitään ja klusteri yrittää ottaa imagen käyttöön. Tässä menee jonkin aikaa. Sitten konsoliin tulee seuraava virheilmoitus:

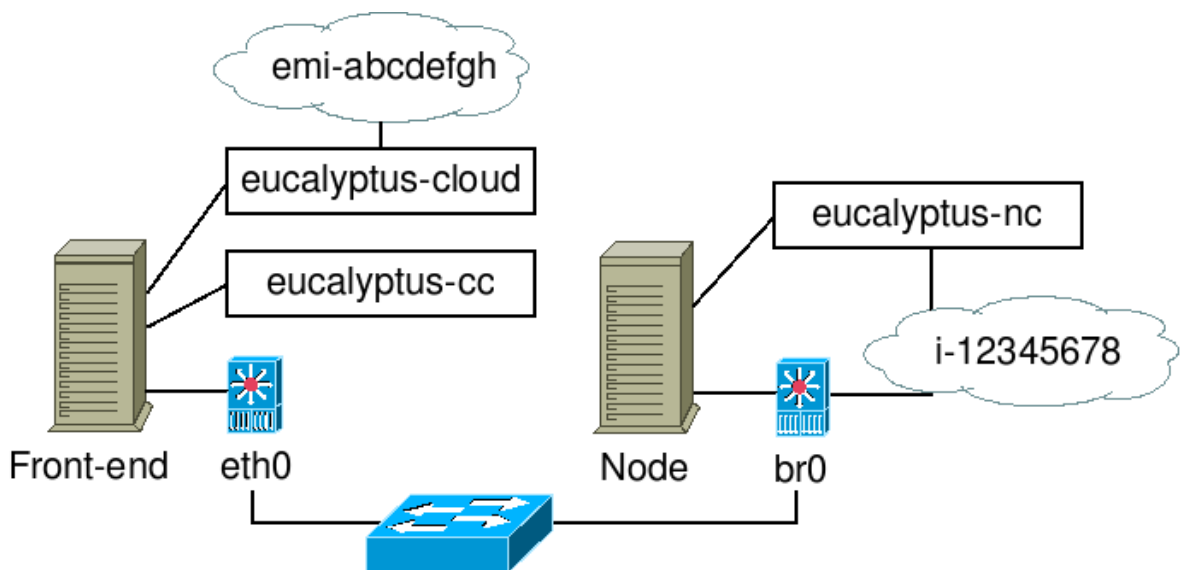
```
klusteri@ykay-c23501:~$ euca-describe-instances
Warning: failed to parse error message from AWS: error in
processing external entity reference
Traceback (most recent call last):
  File "/usr/bin/euca-describe-instances", line 137, in
<module>
    main()
  File "/usr/bin/euca-describe-instances", line 133, in main
    euca.display_error_and_exit('%s' % ex)
                                File
                                "/usr/lib/python2.6/dist-
packages/euca2ools/__init__.py",      line      996,      in
display_error_and_exit
    dom = minidom.parseString(msg)
  File "/usr/lib/python2.6/xml/dom/minidom.py", line 1928,
in parseString
    return expatbuilder.parseString(string)
  File "/usr/lib/python2.6/xml/dom/expatbuilder.py", line
940, in parseString
    return builder.parseString(string)
  File "/usr/lib/python2.6/xml/dom/expatbuilder.py", line
223, in parseString
    parser.Parse(string, True)
```

`xml.parsers.expat.ExpatError: mismatched tag: line 1, column 376`

Tähän ongelmaan yritettiin kauan aikaa etsiä ratkaisua, mutta ongelmaa ei saatu ratkaistua.

## 5 TULOKSET

Klusteria ei saatu lopullisesti toimimaan kunnolla erinäisistä syistä. Työssä käytettävien koneiden kokoonpanot eivät olleet riittävät. Välityspalvelin sotki myös työn testausta paljon. Palaamme näihin ongelmiin kappaleessa 6. Ubuntu Enterprise Cloud ratkaisu on kuitenkin helppokäyttöisin ja vaivattomin klusterointiratkaisu Linuxille.



KUVIO 12. Malli rakennetusta klusterista (Kirkland 2010.)

Kun klusterin rakentaa annettujen ohjeiden mukaan, pitäisi sen näyttää kuten kuviossa 12. Kuitenkaan rakentamamme klusteri ei aivan täysin vastaa mallikuvaa. Ainut minkä saatiin toimimaan oli web-serveri, jossa on testi HTML-sivu.



## It works!

This is the default web page for this server.

The web server software is running but no content has been added, yet.



## KUVIO 13. Web-serverin testisivu

Web-serverille pääsee serverikoneen julkisella IP-osoitteella, eli 10.65.0.90. Myös samalla osoitteella pääsemme samalle konfigurointisivulle kuin mitä menimme aikaisemmin osoitteella 192.168.1.101. Selaimen osoitekenttään täytyy vain muistaa laittaa "https://" eikä "http://" ja lisäksi vielä portti, joka on 8443.

Klusterin hallintasivua voi siis konfiguroida julkisesti mistä vain, kunhan serverikone ja noodikone ovat käynnissä eli klusteri on periaatteessa toiminnassa. Tässä kohdassa täytyy muistaa, että julkinen konfigurointisivu ei toimi, jos välityspalvelin on käytössä.

## 6 JOHTOPÄÄTÖKSET

Työssä rakennettua klusteria ei saatu kunnolla toimimaan muutamasta syystä. Ensinnäkin tietokoneiden kokoonpanot eivät olleet aivan riittävät, koska koneiden prosessorit eivät tukeneet virtualisointia, joka mainitaan minimivaatimuksissa. Koneissa oli myös hieman liian pienet kiintolevyt klusterin rakentamiseen. Yksi iso ongelma-kohta oli myös välityspalvelin, joka sekoitti nettiliikennettä. UEC on rakennettu niin, että nettiyhteyttä tarvitaan, jotta klusteri saadaan kunnolla toimimaan ja tarvittavat puuttuvat paketit voidaan hakea netistä. Kun rakennetaan UEC:llä klusteria, paras vaihtoehto nettiyhteyden luomiseen on suora nettiyhteys.

TAULUKKO 1. Suositukset klusteriserverille (Kirkland 2010.)

Hardware	Minimum	Suggested	Notes
CPU	1GHz	2 x 2GHz	for an all-in-one front end, it helps to have at least a dual core processor
Memory	2GB	4GB	the Java web front end benefits from lots of available memory
Disk	5400rpm IDE	7200rpm SATA	slower disks will work, but will yield much longer instance startup times
Disk Space	40GB	200GB	40GB is only enough space for only a single image, cache, etc., Eucalyptus does <i>not</i> like to run out of disk space
Networking	100Mbps	1000Mbps	machine images are hundreds of MB, and need to be copied over the network to nodes

TAULUKKO 2. Suositukset noodille (Kirkland 2010.)

Hardware	Minimum	Suggested	Notes
CPU	VT extensions	VT, 64-bit, Multicore	64-bit can run both i386, and amd64 instances; by default, Eucalyptus will only run 1 VM per CPU core on a Node
Memory	1GB	4GB	additional memory means more, and larger guests
Disk	5400rpm IDE	7200rpm SATA or SCSI	Eucalyptus nodes are disk-intensive; I/O wait will likely be the performance bottleneck
Disk Space	40GB	100GB	images will be cached locally, Eucalyptus does <i>not</i> like to run out of disk space
Networking	100Mbps	1000Mbps	machine images are hundreds of MB, and need to be copied over the network to nodes

Kun vertaamme molempia taulukoita työssä käytettyjen tietokoneiden kokoonpanoihin, niin huomataan, että koneiden kokoonpanot riittivät melko niukasti minimivaatimuksiin. Prosessoreiden osalta minimivaatimukset eivät täyttyneet.

UEC:llä rakennettu klusteri pitäisi olla aikalailla helpoin ja vaivattomin ratkaisu Linuxille, kunhan vain poistamme nuo muutamat ongelmakohdat.

Työn toteuttamiseen meni melko paljon aikaa ja suurin osa ajasta meni työn alkuvaiheessa hieman hukkaan. Aluksi työssä oli tarkoitus tutkia klusterin rakentamista Matlab ohjelmiston ympärille. Työssä käytetty Ubuntun 9.10 versio oli kuitenkin liian uusi, koska Matlabin versio oli sen verran vanha, ettei sitä saanut toimimaan. Samoin työssä oli tarkoitus rakentaa klusteri OSCAR-nimisellä klusterointiohjelmalla, joka myös olisi tarvinnut vanhemman Ubuntun version (8.04 LTS), jotta

ohjelman olisi saanut toimimaan. Näiden kahden ohjelman kanssa aikaa valui hukkaan, kunnes todettiin, että Ubuntun Server käyttöjärjestelmässä on valmiit työkalut klusterin rakentamiseen. Aika siis loppui hieman kesken ja itse klusterointia ei päässyt kunnolla testaamaan.

Työn tutkimiseen täytyi lukea melko paljon englanninkielistä materiaalia ja Linux käyttöjärjestelmä tuli tutuksi työtä tehtäessä. Valitettavaa on kuitenkin, että klusteroinnin ja pilvipalvelimen rakentamiseen ei ole kovin paljoa materiaalia. UEC on kuitenkin tällä hetkellä jo melko hyvä valmis paketti Linux klusterin rakentamiseen ja sitä kehitetään kokoajan eteenpäin. Kun mietitään hyviä ilmaisia ratkaisuja, niin UEC on siihen erinomainen.



## LÄHTEET

### Sähköiset lähteet

Creative Commons 3.0 2010. Www-dokumentti. Saatavissa:  
<http://linux.fi/wiki/Etusivu>. Luettu 5.1.2010.

Free Software Foundation, Inc. 2010. Www-dokumentti. Saatavissa:  
<http://www.gnu.org/>. Luettu 5.1.2010.

Squidoo, LLC 2010. Www-dokumentti. Saatavissa:  
<http://www.squidoo.com/linux-clustering>. Luettu 16.2.2010.

Creative Commons Attribution/Share-Alike 2010. Www-dokumentti.  
Saatavissa:  
[http://en.wikipedia.org/wiki/Cluster\\_\(computing\)](http://en.wikipedia.org/wiki/Cluster_(computing)). Luettu 16.2.2010.

Canonical Ltd 2010. Www-dokumentti. Saatavissa:  
<http://www.ubuntu.com/cloud/private-steps>. Luettu 23.3.2010

Kirkland, D. 2010. Creative Commons Attribution-ShareAlike 3.0. Www-dokumentti.  
Saatavissa: <https://help.ubuntu.com/community/UEC/NodeInstallation>.  
Luettu 24.3.2010.

Kirkland, D. 2010. Creative Commons Attribution-ShareAlike 3.0. Www-dokumentti.  
Saatavissa: <https://help.ubuntu.com/community/UEC/CDInstall>.  
Luettu 24.3.2010.

Kirkland, D. 2010. Creative Commons Attribution-ShareAlike 3.0. Www-dokumentti.

Saatavissa:

<https://help.ubuntu.com/community/UEC/PackageInstall9.10>.

Luettu 24.3.2010.